Overview of the IPCC WG1 4AR Model Output Database

Robert Drach

Program for Climate Model Diagnosis and Intercomparison

Lawrence Livermore National Laboratory

UCRL-PRES-212663

IPCC 4AR Model Output Database

IPCC (Intergovernmental Panel on Climate Change) Fourth Assessment Report will be released in 2007

In September, 2003, PCMDI was asked to host the climate model database for the IPCC Working Group I (WG1).

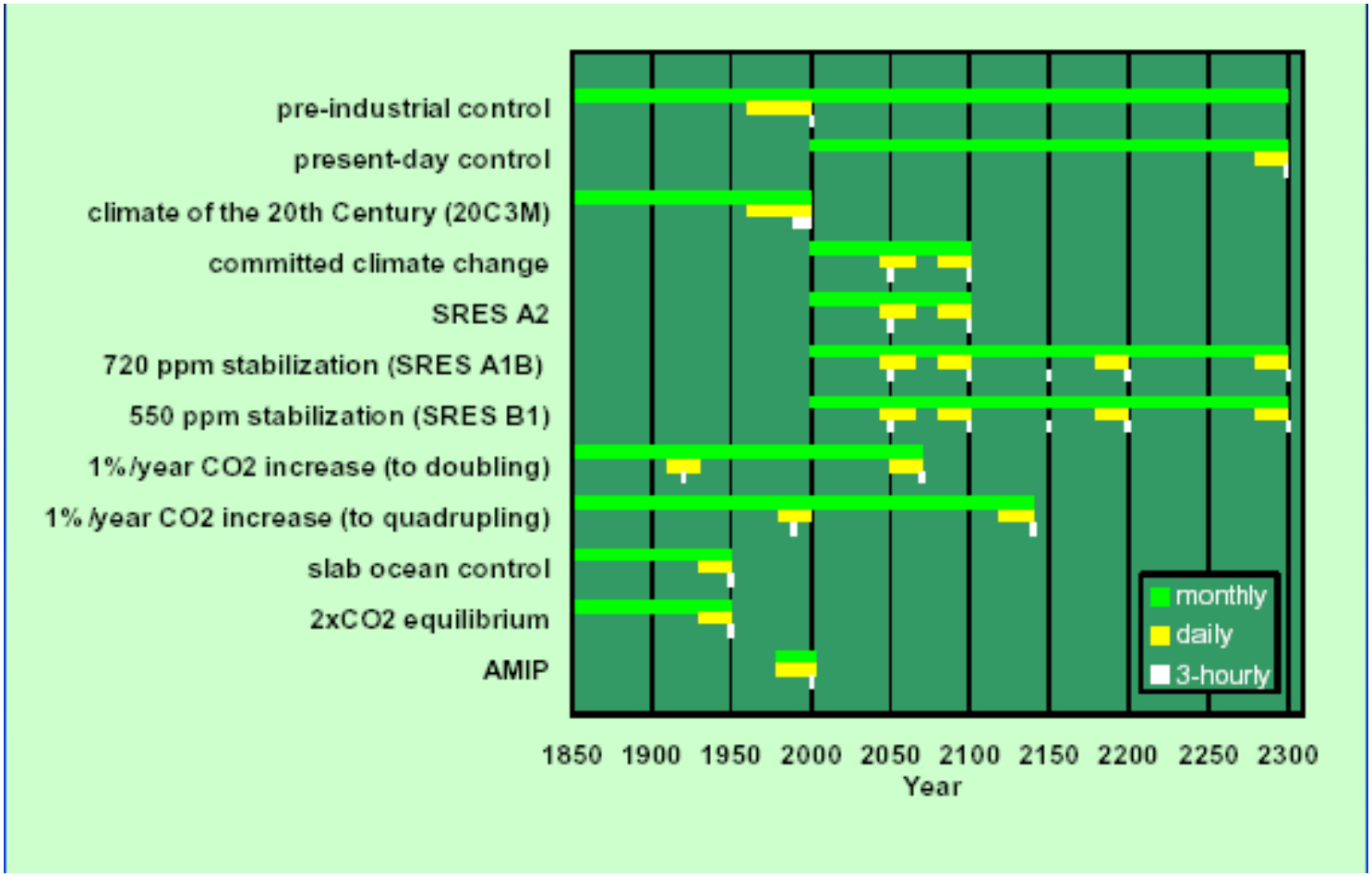
- ***WG1 focuses on the physical climate system: atmosphere, land surface, ocean, and sea ice***
- ***Data from ~20 coupled ocean-atmosphere models***

WG1 specified the 12 scenarios to be run, and the data to be submitted.

Table of Experiments.

	Experiment Name	**Monthly Data and Yearly Data (Extreme Indices)** (submit for each member of ensemble)	**Daily Data** (temperature and precipitation data should be submitted for each member of ensemble, but all other fields should be submitted for only a single ensemble member)	**3-Hourly Data** (submit for a single ensemble member)	**Notes**
1	pre-industrial control experiment	> 100 years (~500 years)	40 years that can best be compared to years 1961-2000 (i.e., through the end of year 2000) of the 20C3M expt.	last year of reported daily data (i.e., corresponding to year 2000 of the 20C3M expt.)	control for experiments 3-7 and for some models also the control for experiments 8-9. There will be no anthropogenic or natural forcing in this control. The control experiment should be long enough to extend to the furthest point in time reached by the end of the perturbation experiments (which presumably branch from it). Thus the control should allow us to subtract any residual, unforced drift from all perturbation simulations.
					for most models this experiment is not needed, but for

Scenarios


pre-industrial control
present-day control
climate of the 20th Century (20C3M)
committed climate change
SRES A2
720 ppm stabilization (SRES A1B)
550 ppm stabilization (SRES B1)
1%/year CO2 increase (to doubling)
1%/year CO2 increase (to quadrupling)
slab ocean control
2xCO2 equilibrium
AMIP
monthly
daily
3-hourly
1850 1900 1950 2000 2050 2100 2150 2200 2250 2300
Year

Models

Modeling Center	Model
BCCR, Norway	BCM 2.0
CCCma, Canada	CGCM3.1
CCSR/NIES/FRCGC (hi-res), Japan	MIROC3.2 (hi-res)
CCSR/NIES/FRCGC (med-res), Japan	MIROC3.2 (med-res)
CNRM, France	CNRM-CM3
CSIRO, Australia	CSIRO Mk3.0
GFDL (CM2.0), USA	GFDL_CM2.0
GFDL (CM2.1), USA	GFDL_CM2.1
GISS (C4x3), USA	C4x3
GISS (Model E-H), USA	Model E-HYCOM
GISS (Model E-R), USA	Model E-Russell

Modeling Center	Model
IAP, China	FGOALS1.0_g
INM, Russia	INMCM3.0
IPSL, France	IPSL-CM4
MPI, Germany	ECHAM 5 / MPI-OM
MRI, Japan	MRI-CGCM2.3.2a
NCAR (CCSM3), USA	CCSM3.0
NCAR (PCM1), USA	PCM1
NCC, China	CSM T63 (Temporal)
UKMO (HadCM3), UK	HadCM3
UKMO (HadGEM1), UK	HadGEM1

IPCC Process

~300 diagnostic subprojects have registered with the IPCC

- **Primary user community in the early stage**

The IPCC process is deadline driven. Among the important dates affecting the Model Output database:

- **Fall, 2004 - Modeling groups begin submission of data**
- **March, 2005 - Workshop on Analyses of Climate Model Simulations for the IPCC AR4. Preliminary results from diagnostic subprojects.**
- **May, 2005 - Papers submitted to peer-reviewed journals, available to IPCC lead authors**
- **December, 2005 - Papers accepted for publication in a peer-reviewed journal**

IPCC WG1 4AR Database

The IPCC deadlines imposed time constraints for development of the Model Output database:

- **Fall, 2004 - Initial data submissions, data organized, cataloged, available for download**
- **Spring, 2005 - Populated database, up to 40TB capacity**

The nature of the data, IPCC process, PCMDI experience set technical requirements:

- **Data to be organized one variable per file**
- **Multifile download capability (e.g., ftp mget)**
- **CF metadata standard conformance + additional constraints**
- **Registration, approval of users by IPCC**
- **Search by variable, scenario, modeling group, temporal frequency**

January, 2004

- **Purchased 40TB RAID server**
- **Began port of Earth System Grid, adaptation for IPCC**

Model Output: Metadata

CF compliance

- **Time encoding, calendars, standard name, grid/axis boundaries…**

Identification of datasets

- **Scenario: 12 defined experiments**
- **Submodel: Atmosphere, ocean, land surface, sea ice**
- **Frequency: 3-hourly, daily, monthly, yearly, fixed**
- **Run number: Ensembles**
- **Model Identification: 22 modeling groups**
- **Variables: Short name, long name**

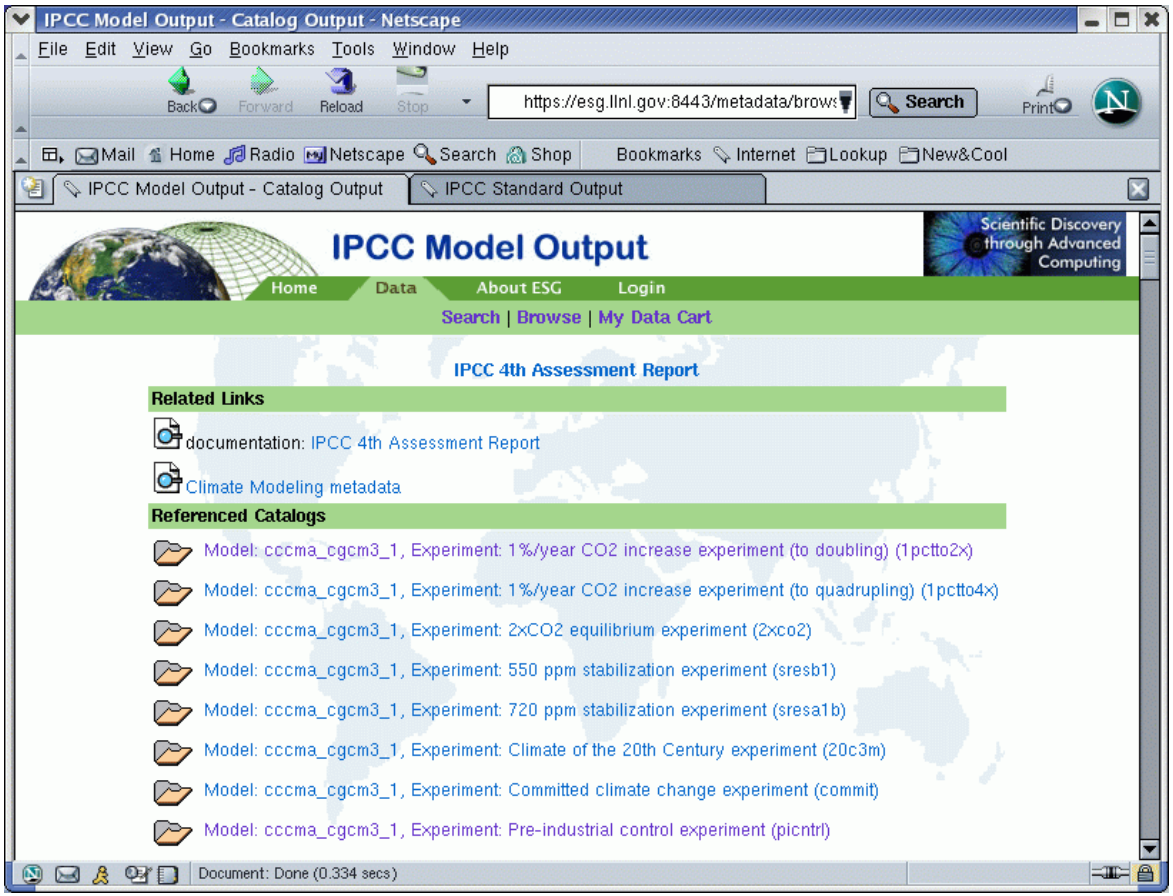
Variable/Frequency/Submodel grouped into Tables:

- **A1: monthly atmospheric**
- **O1: monthly ocean**
- **A2: daily atmospheric**
- **…**

Data can be viewed in different ways

Portal:

- **Model -> Scenario -> Run -> Frequency -> Variable/Submodel ->Files**



Views of data

FTP

- **Scenario -> Submodel -> Frequency -> Variable -> Model -> Run -> Files**

Search by Variable

QUERY: scenario=any, model=cccma_cgcm3_1, frequency=monthly, variable=temperature

Download Files

To download files, select them and then proceed to download.

1-50 of 161 datafiles | next >

Datafile	Metadata	Format	Type	Size	Download
1. pcmdi.ipcc4.cccma_cgcm3_1.1pctto2x.run1.monthly.tas_a1_1pctto2x_1_cgcm3.1_t47_1850_2069.nc	Surface Air Temperature	NetCDF	gridded	48730096	LLNL HRM
2. pcmdi.ipcc4.cccma_cgcm3_1.1pctto2x.run1.monthly.tas_a1_1pctto2x_1_cgcm3.1_t63_1850_2069.nc	Surface Air Temperature	NetCDF	gridded	86578128	LLNL HRM
3. pcmdi.ipcc4.cccma_cgcm3_1.1pctto2x.run1.monthly.ta_a1_1pctto2x_1_cgcm3.1_t47_1850_2069.nc	Temperature	NetCDF	gridded	827297716	LLNL HRM
4. pcmdi.ipcc4.cccma_cgcm3_1.1pctto2x.run1.monthly.ta_a1_1pctto2x_1_cgcm3.1_t63_1850_2069.nc	Temperature	NetCDF	gridded	1470698544	LLNL HRM
5. pcmdi.ipcc4.cccma_cgcm3_1.1pctto2x.run1.monthly.thetao_o1_1pctto2x_1_cgcm3.1_t47_1850_1919.nc	Potential Temperature	NetCDF	gridded	1796044224	LLNL HRM
6. pcmdi.ipcc4.cccma_cgcm3_1.1pctto2x.run1.monthly.thetao_o1_1pctto2x_1_cgcm3.1_t47_1920_1969.nc	Potential Temperature	NetCDF	gridded	1282891584	LLNL HRM
7. pcmdi.ipcc4.cccma_cgcm3_1.1pctto2x.run1.monthly.thetao_o1_1pctto2x_1_cgcm3.1_t47_1970_2019.nc	Potential Temperature	NetCDF	gridded	1282891584	LLNL HRM
8. pcmdi.ipcc4.cccma_cgcm3_1.1pctto2x.run1.monthly.thetao_o1_1pctto2x_1_cgcm3.1_t47_2020_2069.nc	Potential Temperature	NetCDF	gridded	1282891584	LLNL HRM
9. pcmdi.ipcc4.cccma_cgcm3_1.1pctto2x.run1.monthly.thetao_o1_1pctto2x_1_cgcm3.1_t63_1850_1874.nc	Potential Temperature	NetCDF	gridded	1710510624	LLNL HRM
10. pcmdi.ipcc4.cccma_cgcm3_1.1pctto2x.run1.monthly.thetao_o1_1pctto2x_1_cgcm3.1_t63_1875_1899.nc	Potential Temperature	NetCDF	gridded	1710510624	LLNL HRM
11. pcmdi.ipcc4.cccma_cgcm3_1.1pctto2x.run1.monthly.thetao_o1_1pctto2x_1_cgcm3.1_t63_1900_1919.nc	Potential Temperature	NetCDF	gridded	1368411264	LLNL HRM
12. pcmdi.ipcc4.cccma_cgcm3_1.1pctto2x.run1.monthly.thetao_o1_1pctto2x_1_cgcm3.1_t63_1920_1944.nc	Potential Temperature	NetCDF	gridded	1710510624	LLNL HRM
13. pcmdi.ipcc4.cccma_cgcm3_1.1pctto2x.run1.monthly.thetao_o1_1pctto2x_1_cgcm3.1_t63_1945_1969.nc	Potential Temperature	NetCDF	gridded	1710510624	LLNL HRM

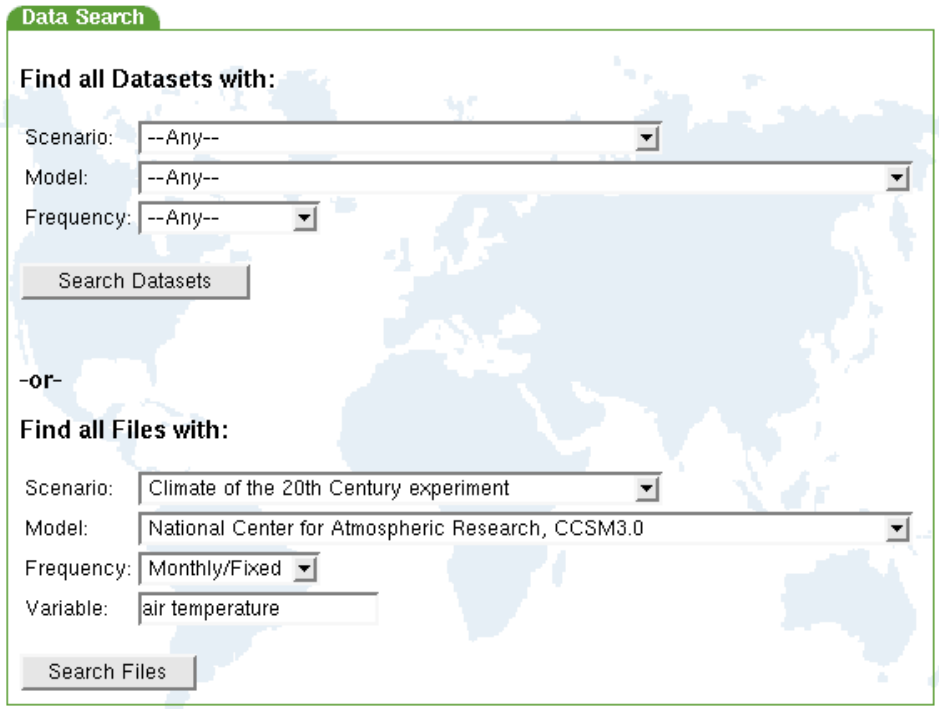
Search

ESG search capability extended to support search on:

- **Model**
- **Scenario**
- **Frequency**
- **Variable**

and return:

- **Datasets**
- **Files**



ESG-II Data Portal was adapted for Model Output database

Summer, 2004 - ESG-II data portal transferred to PCMDI, adapted to IPCC requirements

- **Metadata search extended for scenario, variable, …**
- **Portal registration enables access to portal and FTP server**

November, 2004 - Opened database for registration, downloading

- **Most data transferred by 1TB disk**
- **CMOR software (Doutriaux, Taylor, Peterschmitt) used by most groups to comply with standards**

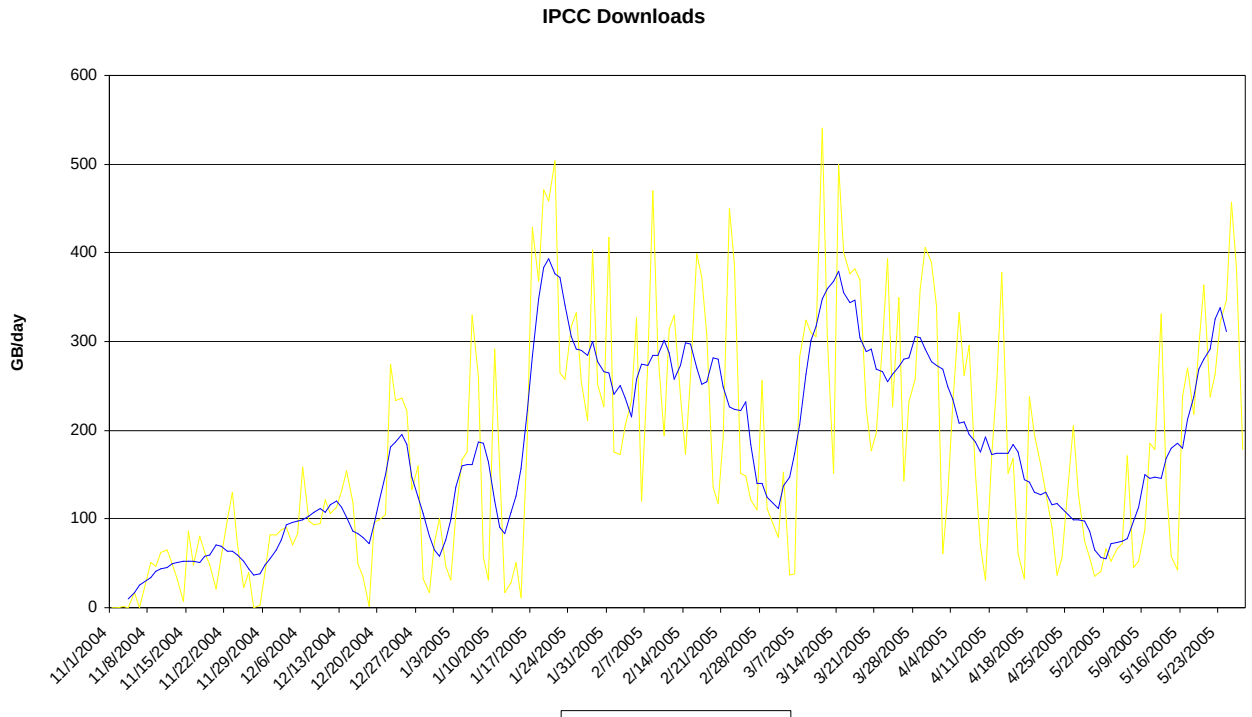
What we expected:

- **~100 users**
- **Modeling groups would send data on one or two disks**
- **Relatively static database - some updates, corrections**

Production period (Nov 2004 - Present)

What actually happened (5/27/05):

- **~350 users**
- **37TB downloaded in 192,000 files**
- **~23TB cataloged, online, 50,000 files**
- **Relatively dynamic database - *many* corrections, resubmissions**



Errata

Errata page details problems with data, resubmission of files. About 70 entries to date.

Errata

Date	Model(s)	Files	Description	Status
11/10/04	ncar_pcm1	cl_A1.20C3M_2.PCM1.atmm.1940-01_cat_1949-12.nc tro3_A1.SRESA1B_2.PCM1.atmm.2010-01_cat_2019-12.nc	Time array is invalid, has zero values at the tail end.	12/10/04 New data files are available.
11/22/04	miroc3_2_hires	/ipcc/[20c3m|1pctto2x|picntrl]/ocn/mo/zos/miroc3_2_hires/run1/zos_O1.nc	Data were withdrawn.	1/10/05 New data files are available.
11/23/04	giss_model_e_r, giss_model_e_h	All rlds_A1, rlus_A1, rldscs_A1, rsutcs_A1 files	Files are wrong.	1/3/05 Replacement files available for rldscs, rsutcs only.
11/29/04	miroc3_2_hires	Scenarios: 20c3m, 1pctto2x, picntrl Variables: - rhopoto - stfmmc - thetao - uo - vo	Data were withdrawn.	1/10/05 New data files are available.
12/06/04	giss_model_e_r	pr_A1.GISS1.1%to2x.nc tas_A1.GISS1.1%to4.nc pr_A1.GISS1.SRESB1.nc pr_A1.GISS1.SRESA1B.run1.nc pr_A1.GISS1.SRESA1B.run3.nc pr_A1.GISS1.SRESA1B.run5.nc	Time values are missing (0-length vector)	1/3/05 Replacement files available for first three files only.
		/ipcc/20c3m/land/mo/snd/miroc3_2_medres/run1/snd_A1.nc /ipcc/20c3m/land/mo/snd/miroc3_2_medres/run2/snd_A1.nc /ipcc/20c3m/land/mo/snd/miroc3_2_medres/run3/snd_A1.nc		

Future work

Adapt current ESG-II work on aggregation server

- **In testing phase**

Expect database to be a valuable resource for several years to come

- **Enlarge user group**
- **Continue to accept new data, data corrections**

Expect volume of downloads to level off after Summer, 2005